Chapter 1

Pulse Modulation Systems

1.1 Introduction

Many signals are inherently digital in nature such as telegraph, computer outputs, pulsed radar, and sonar signals. Speech, TV, facsimile, and telemeter data signals are all examples of analogue signals. The current widespread use of digital signaling is the result of many factors:

- The relative simplicity of digital circuit design and the ease with which one can apply integrated circuit techniques to digital circuitry.

- The ever-increasing use and availability of digital signal processing techniques.
- Widespread use of computers in handling all kinds of data.
- The ability of digital signals to be coded to minimize the effects of noise and interference.

The principle of sampling can be explained using Fig.1.1. The continuous signal $f(t)$ is multiplied by the impulse train $s(t)$ of (sampling signal) to produce $f_s(t)$, the discrete or the digital signal.

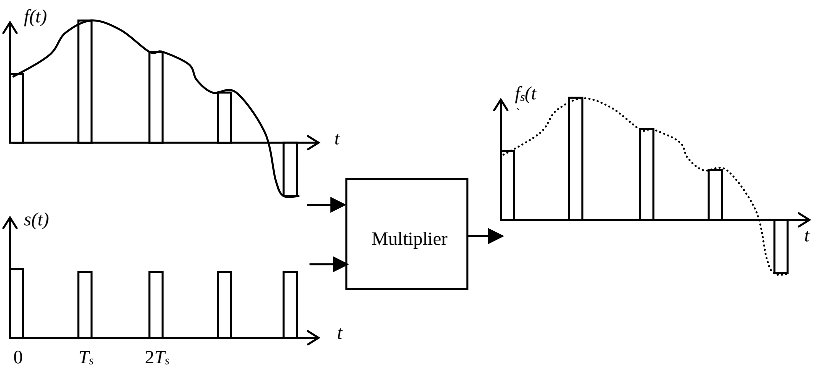

Fig.1.1: Principle of Sampling

The sampling signal $s(t)$ is periodic, and can be represented by Fourier expansion to be:

$$s(t) = \frac{\tau}{T_s} + \frac{2\tau}{T_s}\left(\cos\omega_s t + \cos 2\omega_s t + \ldots\right) \qquad (1.1)$$

$$s(t) = \frac{\tau}{T_s} + \frac{2\tau}{T_s}\left(\cos 2\pi \frac{t}{T_s} + \cos 4\pi \frac{t}{T_s} + \ldots\right) \qquad (1.2)$$

T_s is called the sampling time. [Exercise 1.1: Prove the equation 1.1?]. Assume $f(t)$ is a base band signal of maximum frequency f_m. When $T_s = 1/2f_m$, the product $f(t)\ s(t)$ is given:

$$f(t)s(t) =$$

$$\frac{\tau}{T_s} f(t) + \frac{\tau}{T_s}\left[2f(t)\cos 2\pi(2f_m)t + 2f(t)\cos 2\pi(4f_m)t..\right] \qquad (1.3)$$

The first term is the signal $f(t)$ itself, the second term gives rise to the double-sideband suppressed-carrier with carrier frequency $2f_m$, the third is at $4f_m$, etc. Using a low pass filter with sharp cutoff, the original continuous signal $f(t)$ can be recovered exactly from the discrete digital signal $f_s(t)$ conditioned that:

$$T_s < 1/2 f_m \qquad (1.4)$$

Fig.1.2 shows the spectrum of the sampler product $f(t)\ s(t)$ with impulse train $s(t)$ when the spectrum of the analogue signal $f(t)$ is assumed $F(s)$ as indicated in Figure.

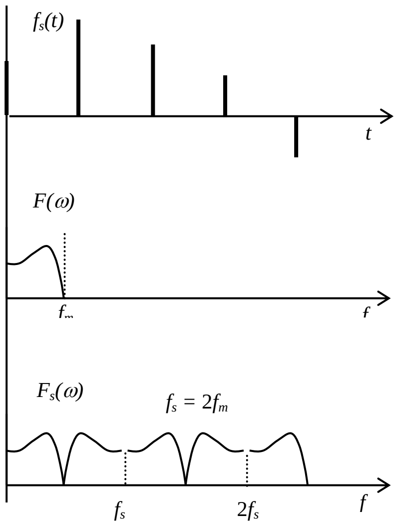

Fig.1.2: Spectrum of Product

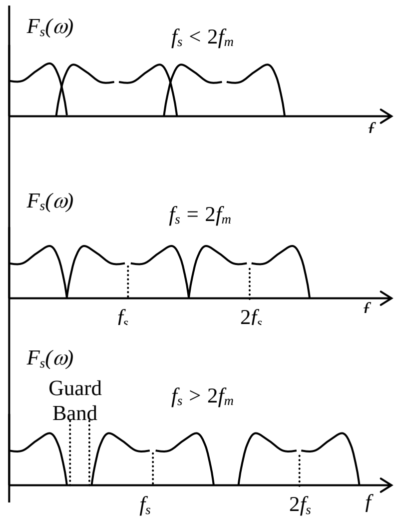

Fig.1.3: Effect of Sampling

1.2 Sampling Rate and Nyquist Sampling

There is a relation between the rate at which a signal varies and the number of pulses needed to reproduce it exactly. The rate at which a signal varies is related to its maximum frequency component or bandwidth. Sampling theorem tells that at least $2B$ uniformly spaced samples are needed every second in order to reproduce the signal without distortion.

Fig.1.3 indicates the spectrum of the product when the sampling frequency f_s is less than, equal, or greater than the twice of the maximum frequency f_m of the signal $f(t)$. The minimum allowable sampling rate is the Nyquist sampling for which $f_s = 2f_m$ or $2B$. However, as in Fig.1.3.b, it requires a filter with infinitely sharp cutoff that is not realizable in practice.

The case for which $f_s > 2f_m$ is shown in Fig.1.3.c. There is a gap between the upper limit of the spectrum of the baseband signal and the lower limit of the DSB-SC spectrum. This is called a guard band is always required in practice. The filter used to select the signal $f(t)$ need not have an infinitely sharp cutoff.

For the case where $f_s < 2f_m$ shown in Fig.1.3.a, there is an overlapping between the spectrum of $f(t)$ itself and the spectrum of the DSB.SC signal. Accordingly, no filtering operation will allow an exact recovery of the original signal. The phenomenon of overlapping and the distortion that results is termed ***aliasing***. For too low sampling rate the signal may change radically between sampling times. We thus lose information and produce a distorted output as shown in Fig.1.4. In order not to lose the signal dips and rises, additional sampling pulses must be added as shown.

As a summary, the increase in sampling above the Nyquist rate increases the guard band, thereby easing filtering. On the other hand, this extends the bandwidth required for transmitting the sampled signal so that compromise should be made.

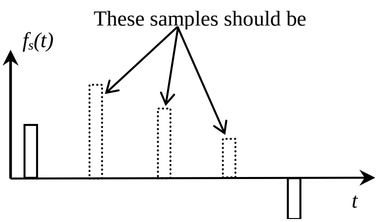

Fig.1.4: Sampling Frequency Too Low

That is to sample at a somewhat higher rate to ensure separation of the spectra and to simplify filtering. As an example, speech transmitted via telephone is generally filtered to bandwidth B =3.3 KHz. The Nyquist rate is thus 6.6 KHz. However, for digital transmission, the speech is normally sampled at an 8 KHz rate.

1.3 Pulse Amplitude Modulation PAM

Fig.1.5 illustrating how the sampling principle may be used to transmit a number of band limited signals over a single communication channel. As the rotary arm of the switch swings around, it samples each signal sequentially. The rotary switch at the receiving end is in synchronism with the switch at the sending end. With each revolution, one sample is taken of each input signal. The switches must make at least $2f_m$ revolutions per second. When the train of pulses corresponding to the samples of each signal are modulated in amplitude, the scheme of sampling is called pulse amplitude modulation PAM as indicated in Fig.1.5 on multiplexing two signals only.

Suppose N independent baseband signals, each of which is band limited to f_m, are transmitted simultaneously using PAM time division multiplexing. At least, the channel need not have a bandwidth larger than Nf_m. So, multiplexing N signals by PAM requires no more bandwidth than would be required to multiplex these signals by FDM using SSB transmission.

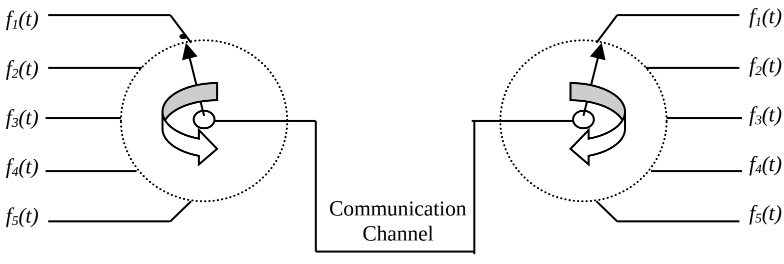

Fig.1.5: Time Division Multiplexing and Sampling Principle

1.3.1 Instantaneous Sampling

The instantaneous sampling is achieved when the sampling pulse train width is very small. Such instantaneous sampling is hardly feasible.

Instantaneous samples at the transmitting end of the channel have infinitesimal energy, and when transmitted through a band limited channel give rise to signals having a peak value which is infinitesimally small. Such signals will be lost in background noise, Fig.1.6.

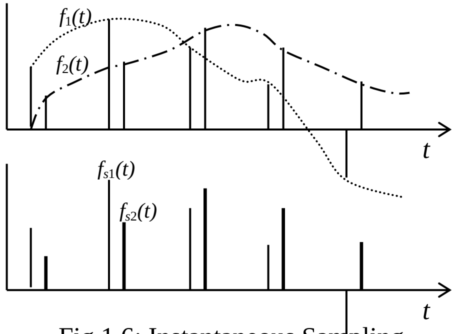

Fig.1.6: Instantaneous Sampling

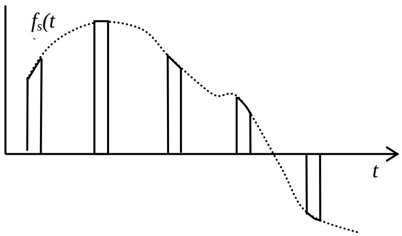

Fig.1.7: PAM (Natural Sampling)

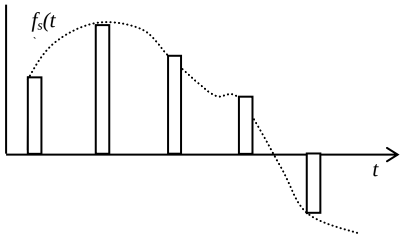

Fig.1.8: PAM (Flat-top Sampling)

1.3.2 Natural Sampling

Here, as in Fig.1.7, the train pulses have duration τ and is separated by the sampling time T_s. The top of the samples are not flat. With samples of duration τ, it is not possible to completely eliminate crosstalk generated in the channel. If N signals are multiplexed, the maximum sample duration τ is T_s/N. It is advantageous to make τ as large as possible to increase the output. However, to suppress crosstalk τ should be minimized.

1.3.3 Flat-top Sampling

In sampling of this type, Fig.1.8, the baseband signal $f(t)$ cannot be recovered exactly by simply passing the samples through an ideal low pass filter. However, the distortion need not be large. Flat-top sampling has the merit that it simplifies the design.

1.4 Signal Recovery through Holding

Assume an N signals are being time division multiplexed. As N increases, maximum ratio of sample duration to the sampling interval $t/Ts = 1/N$ becomes smaller and so does the output signal. There is an alternative method of recovery which rises the level of the output signal without the use of amplifiers which may introduce noise. This method has the advantage of filtering being quite adequate, but has the disadvantage that some distortion must be accepted. The method is illustrated in Fig.1.9 where the sample pulses are extended or hold until the occurrence of the next sample.

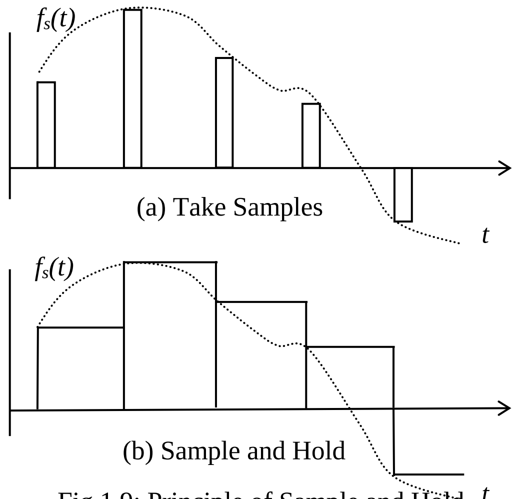

Fig.1.9: Principle of Sample and Hold

1.5 Crosstalk for Communication Channel

Reason: In general, the communication channel has got or has been given a limited bandwidth. However, the spectrum of the sampled signal extends to infinity. This introduces crosstalk.

How to avoid? In principle, crosstalk may be avoided by instantaneous sampling, i.e. instantaneous commutation and de-commutation).

1.5.1 Crosstalk for High Frequency Cut-off

Definition: Crosstalk may result with finite-duration sampling as a result of the upper frequency cut-off of the channel.

Modelling: Assume that an RC circuit (LPF) represents the channel. Its upper frequency 3dB cut-off will be:

$$f_{High-Frequency-Cutoff} = f_{LPF} = \frac{1}{2\pi RC} \tag{1.5}$$

Significance: Fig.1.10 shows that the crosstalk due to high frequency cut-off is significance only between adjacent time slots (i.e. adjacent channels users).

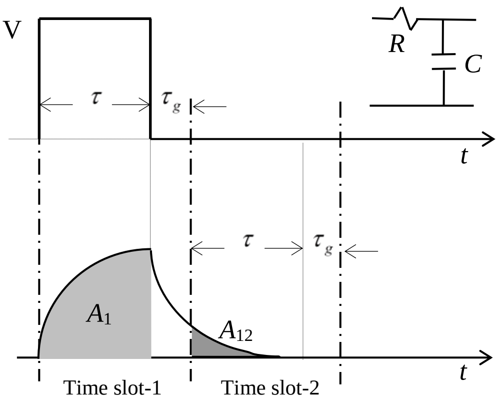

Fig.1.10: Crosstalk of Upper Cut-off (LPF)

Minimizing: To keep crosstalk small, the channel time constant $\tau_c = RC$ must be very small compared to the guard interval $\tau_c << \tau_g$.

Usually: $\tau_g < \tau$ so that $\tau >> \tau_c$.

Crosstalk Factor: If channel 2 is empty and the input pulse is of value V then the crosstalk factor is given by:

$$A_1 \cong \tau V \tag{1.6}$$

$$A_{12} \cong \tau_c V e^{-\tau_g/\tau_c} \tag{1.7}$$

$$\mathrm{K} = \frac{A_{12}}{A_1} = \frac{\tau_c}{\tau} e^{-\tau_g/\tau_c} \tag{1.8}$$

1.5.2 Crosstalk for Low-Frequency Cut-off

Modelling: Here, RC high-pass filter circuit represents the channel. Sample pulse in TS-1 develops an exponential tilt Δ in TS-2 as indicated in Fig.1.11 given as:

$$tilt = \Delta = V\left(1 - e^{-\tau/\tau_c}\right) \tag{1.9}$$

Minimizing: When $\tau << \tau_c$ the exponential tilt is very linear so that crosstalk may be reduced. In such a case Δ approximated as:

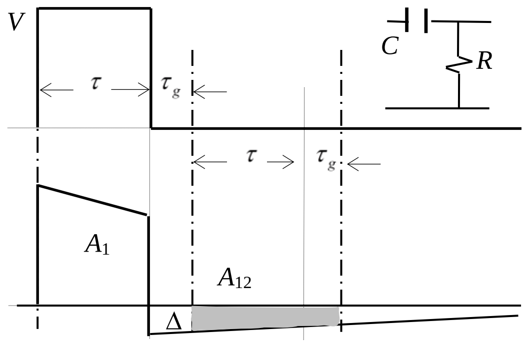

Fig.1.11: Crosstalk of Lower Cut-off (HPF)

$$\Delta = \frac{\tau}{\tau_c} V \tag{1.10}$$

When $\tau_g << \tau_c$, as is certainly the case, so:

$$A_1 \cong \tau V \tag{1.11}$$

$$A_{12} = \tau \Delta \cong \frac{\tau^2 V}{\tau_c} \tag{1.12}$$

Crosstalk Factor: hence:

$$K = \frac{\tau}{\tau_c} \tag{1.13}$$

Important Notice: The crosstalk that results from the upper-frequency bandwidth restriction involves only neighbouring channels. The situation is different for crosstalk caused by low-frequency cut-off.

Exercise 1.2: What is the difference between crosstalk of both high and low frequency cut-offs?

Example 1.1:

Assume the multiplexing of 12 speech channels in a PAM system. The crosstalk in channel N+1 is reduced by a factor 1000 below the desired signal. If channels N and N+1 have the same signal level, what must be the channel bandwidth, the lower and the upper frequency cut-offs of the channel?

Answer 1.1:

For speech: $f_m = B$ = 3.3 kHz, so that $2f_m$ = 6.6 kHz,
Sampling frequency f_s = 8 KHz, so that guard band =1.4 KHz

Samples interval: $T_s = \frac{1}{f_s} = \frac{1}{8000} = 125\mu s$

Time slot duration $\tau + \tau_g$ is given by:

$$\tau + \tau_g = \frac{Sample.Intervel}{No.of.Channels} = \frac{125}{12} = 10.4\mu s$$

However, in practice, sample duration τ is usually two third the time slot, so τ = 6.94 μ Sec. This allow a guard:

τ_g = 3.46 μ Sec.

Upper-frequency cut-off:

$$K = \frac{\tau_c}{\tau} e^{-\tau_g/\tau_c}, \qquad K = 0.001 \qquad 0.001 = \frac{\tau_c}{6.94} e^{-3.46/\tau_c}$$

$$\therefore \tau_c \cong 0.76\,\mu Sec \qquad \therefore f_{c(upper\ frequency\ cutoff)} = \frac{1}{2\pi\tau_c} = 210\ KHz$$

Lower-frequency cut-off:

$$K = \frac{\tau}{\tau_c}, \qquad K = 0.001 \qquad 0.001 = \frac{6.94}{\tau_c} \qquad \therefore \tau_c \cong 6940\,\mu Sec$$

$$f_{c(Lower\ frequency\ cutoff)} = \frac{1}{2\pi(6940)} = 22.93x10^{-6}\,\mu Hz \cong 23 Hz$$

Channel Bandwidth:

Channel bandwidth can be estimated as the difference between the two cutoffs. So, B =210 kHz – 23 Hz = 210 kHz.

1.6 Pulse-Time Modulation

An alternative modulation scheme, rather than the amplitude, some feature of the timing of the pulse, which falls within the time slot, could be forced to vary. Such pulse time modulation (PTM) may be accomplished in a number of ways, such as pulse duration modulation (PDM), or pulse position modulation (PPM). PDM and PPM timing are shown in Fig.1.12. Fig.1.13 shows one method that may be used to generate PDM. PPM can be simply obtained from PDM just by triggering on the negative going pulse edges of PDM.

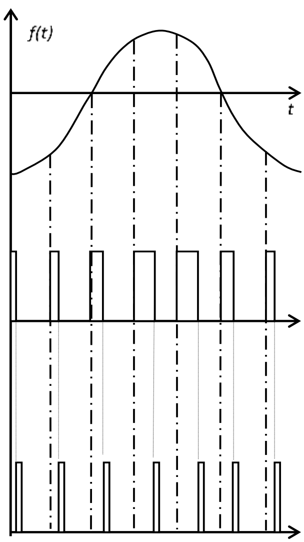

Fig.2.12: PTM Schemes
(a) Baseband signal,
(b) PDM signal,
(c) PPM signal

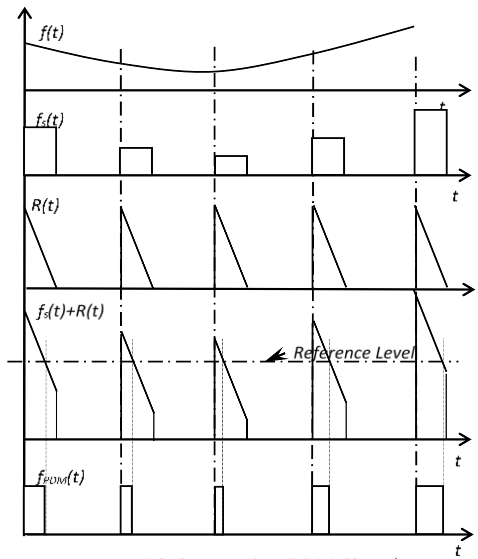

Fig.2.13 Generating PDM Signal
(a) Baseband signal, (d) Sum of a and b,
(b) PAM signal (e) PDM signal
(c) Ramp signal

1.7 Recovery of PDM and PPM

1.7.1 Recovery Using LPF

The spectrum of a naturally sampled PDM waveform with the modulation superimposed on the trailing edge of the pulse is shown

in Fig.1.14 for sinusoidal signal of frequency f_m. It shows how it is possible to recover the signal approximately by LPF limited to f_m.

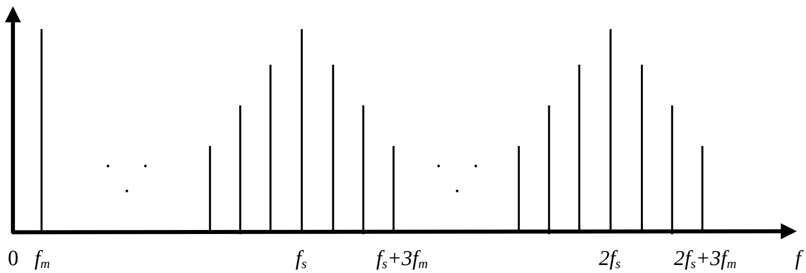

Fig.1.14: Spectrum of PDM signal with Sinusoidal Signal

In PAM if $f_s > 2f_m$, it is possible to adjust the cutoff frequency of the LPF such that it passes only the baseband spectral component. This because each of the lines at f_s, $2f_s$, etc. would be accompanied by a single sideband pair. However, in PDM of Fig.1.14, the sideband spectra extend indefinitely outward.

Therefore, any LPF must include some of the lower sideband components of the carrier at f_s and to the carrier at $2f_s$, etc., causing a distortion. Such distortion can be minimized by raising the sampling frequency and making the variation small compared to the time interval between successive pulses.

1.7.2 Conversion to PAM

Applying inverse operations of generation can make a precise recovery of PDM and PPM. That is to convert PDM or PPM back to PAM and then to demodulate PAM using the appropriate LPF filters. This technique is illustrated in Fig.1.15 for PDM.

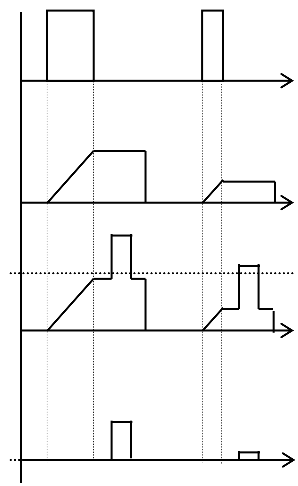

Fig.1.15

1.8 Synchronization

In time division multiplexing TDM, the commutation at the transmitting end and de-commutation at the receiving end must be synchronized with each other. So, a clock source is used at the transmitter to switch the commutator from one slot to the next. It may be a sinusoidal oscillator in addition to some wave shaping circuits. Its repetition frequency equals the product of the sampling frequency and the number of channels.

On the other hand, at receiver end, a clock signal is required to keep the de-commutator running at the same rate of transmitter clock. To keep the proper rate and the channel selection at the two ends, it is necessary to transmit information to tell the demodulator in each cycle of operation the exact timing.

Usually, only one time slot (or more as in several thousand telephone multiplex system) is used per frame for the purpose of transmitting the synchronization information.

On reception, the synchronization slot could be distinguished by its nature such as the amplitude of the PAM signal.

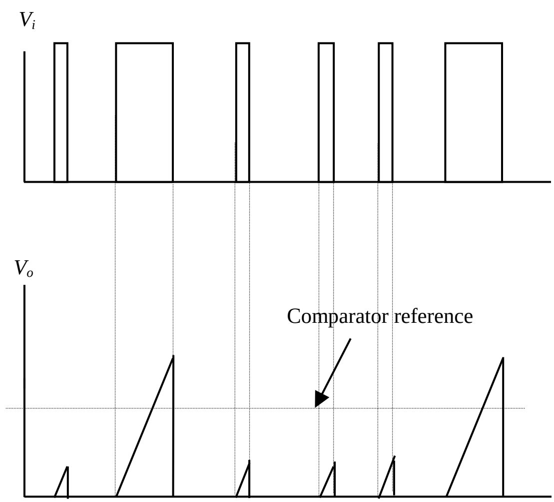

Fig.1.16: Synchronization Circuit for PPM

Fig.1.16 illustrates the case of synchronization pulses of three channels multiplexed using PPM pulse modulation formats.

1.9 Communication Channel Noise

Noise Causes: If a signal is transmitted by radio, when it arrives at its destination, it will be attenuated and combined with noise due to all manners of random electrical disturbances during propagation.

Even wire transmission suffers attenuation, distortion, and noise by an amount increases with path length. Unless coaxial cable is used, electrical noise and crosstalk disturbance will be picked up. Even, low frequency magnetic fields will penetrate the outer conductor of coaxial and induce signals on it.

Reducing Noise: One way of resolving this problem is to raise the signal level. Such a solution is hardly feasible. **The second** method is to use a repeater at the midpoint of long communication. This repeater will raise the signal level and the noise introduced in the first half. So it improves the received signal-to-noise ratio.

Next is to use additional repeaters, but these repeaters as amplifiers produce more noise.

1.10 Quantization of Signals

Once noise has been introduced at any place along the channel, we are stuck with it. This can only be modified by quantization that helps to reconstruct the signal and reject noise completely as soon as the noise is not so severe.

Usefulness of Quantization:

- Quantization helps to reconstruct the signal and reject noise.
- Quantized signal has the great merit that it is separable from additive noise.

Either the sampled pulses may be quantized or both quantization and sampling may be performed simultaneously.

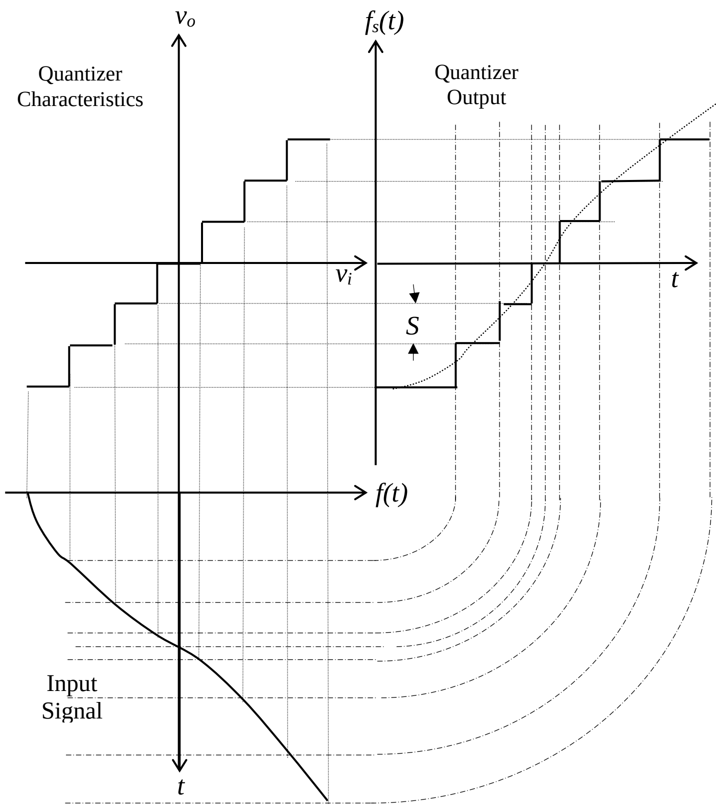

Fig.1.17: Operation of Quantization

[a] Quantizer Input-Output Characteristic Is a Staircase Waveform:

- Quantized output changes abruptly by a quantum jump *S*.
- Quantized signal is an approximation to the original signal.
- The quality of approximation may be improved by reducing the size *S* of the steps (see Fig.1.17).

[b] Practical Levels for Quantization: Tests for TV indicate that:

- 64 levels gives only fairly good color TV performance.
- 512 levels is used to obtain quality of commercial color TV.

Tests for speech indicate that:

- 2 levels are understandable although quite noisy.
- 8 or 16 are sufficient for a good intelligibility of the speech.
- However, 128 or 256 levels are usually used to ensure high quality transmission for digital telephony.

[c] Quantizing Effect:

- In quantizing, some details information will be lost.
- So, it is impossible to reconstruct the original signal exactly from its quantized version.
- However, there is no need to transmit all signal details: **Why?**
    - Our ears in hearing and our eyes watching is limited with regard to gradation of signals it can distinguish.
    - Also, due to noise, the detector will not be able to distinguish fine variation.

[d] Reconstruction and Noise Removal:

When a signal is reconstructed by quantization at the repeater, the added noise is removed. However, sometimes the process of quantization induces a level error as indicated in Fig.1.18. This probability of level error P_q can be reduced by increasing the step size S. Nevertheless, increasing S also increases the discrepancy between $f(t)$ and $f_s(t)$. This difference causes what is referred to as "***quantization error***".

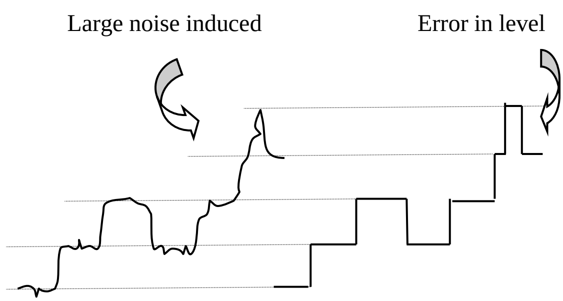

Fig.1.18: Noise Removing or Level Error

1.11 Quantization Error

The difference between the original signal *f(t)* and its quantized approximation signal $f_s(t)$ is called for as the ***quantization error*** as indicated in Fig.1.18. Quantization error takes place on generation of quantized version.

- **Overload distortion (peak limiting):** occurs if the magnitude of the sample exceeds the highest quantization interval.
- **Resolution:** is the minimum voltage that can be decoded by the digital to analogue converter DAC at the receiver. It equals the voltage of the least significant bit.
- **Why quantization noise?**
  - Is equivalent to AWGN since it alters the signal amplitude.
  - It may be added to or subtracted from the actual signal.
  - Its maximum value is 1/2 of the least significant bit voltage:

$$Q_e = \frac{V_{LSB}}{2} \tag{1.14}$$

The mean square quantization error voltage at the quantizer output can be shown to be as follows: [**Exercise 1.3** Prove it?**]**

$$\overline{e^2} = \frac{S^2}{12} \tag{1.15}$$

Meanwhile, the quantization error voltage will be as follows:

$$N_{ov} = \sqrt{\overline{e^2}} = \frac{S}{\sqrt{12}} \tag{1.16}$$

On the other hand, the signal voltage at the quantizer output is:

$$S_{ov} = \frac{SM}{2} \tag{1.17}$$

So, the signal to quantization error voltage ratio will be:

$$\frac{S_{ov}}{N_{ov}} = \sqrt{3}\, M \tag{1.18}$$

Whereas, the signal to quantization error power ratio will be as follows:

$$\frac{S_o}{N_o} = 3M^2 \tag{1.19}$$

By converting to disable it becomes:

$$\left.\frac{S_o}{N_o}\right)_{dB} = 4.8 + 20\log_{10} M \tag{1.20}$$

1.12 Digital Pulse Modulation

Pulse modulation systems can be analogue or digital. In analogue pulse modulation, the amplitude; PAM, width; PDM, or position; PPM of a pulse can vary over a continuous range in accordance with the message amplitude at the sampling instant. In digital pulse modulation, the transmitted samples take on only discrete values such as in delta modulation; DM and pulse code modulation; PCM. Digital modulation has the **advantages**:

- Information is only conveyed be the presence or the absence of a pulse. The pulse shape or exact amplitude value is not significant.
- Signals may be correct re-generated on transmission.
- The using of all digital circuitry.
- It is applicable to all forms of digital signal processing.
- Noise and interference can be minimized through using coding techniques.

1.12.1 Delta Modulation DM

Delta modulation DM is digital modulation technique in which message signal is encoded into sequence of binary symbols.

Description:

- A simple one bit PCM code to achieve digital transmission.
- Indicates whether sample is larger or smaller of previous one.
- Extremely simple modulator and demodulator circuits.

1.12.2 DM Transmitter

Fig.1.19.a illustrates the general block diagram of delta modulator where $f_s(t)$ is a stair step approximation of $f(t)$. The modulator input is:

$$d(t) = f(t) - f_s(t) \tag{1.21}$$

The signal $d(t)$ is hard-limited and multiplied by the pulse generator to yield the output $x_c(t)$ while $f_s(t)$ is by integration:

$$x_c(t) = \Delta(t)\sum_{-\infty}^{\infty} \delta(t - nT_s) = \sum_{-\infty}^{\infty}\Delta(nT_s)\,\delta(t - nT_s) \tag{1.22}$$

$$f_s(t) = \int_0^t x_c(t)dt = \sum_{n=-\infty}^{\infty} \Delta(nT_s)\int_0^t \delta(y - nT_s)dy \tag{1.23}$$

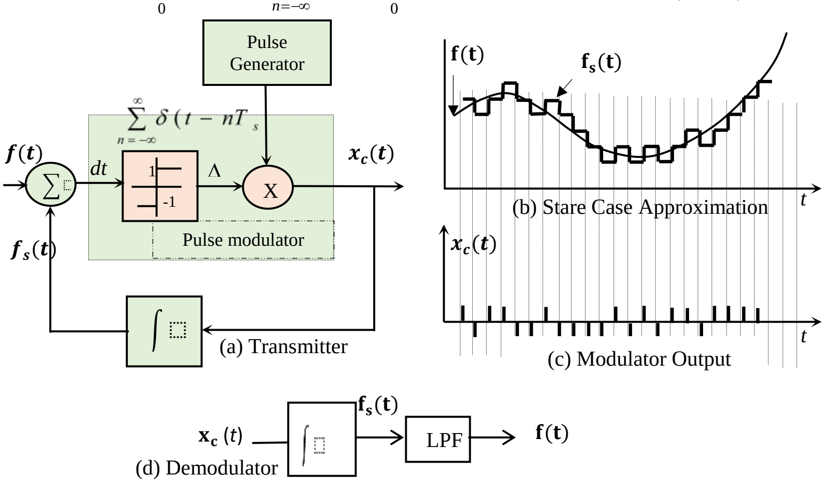

Fig.1.19: Delta Modulation System

DM Demodulator:

Demodulation of DM is simply accomplished by integrating $x_c(t)$ to form the stair step approximation $f_s(t)$ that is low pass filtered to suppress the discrete jumps as indicated in Fig.1.19.d. Since LPF approximates an integrator, it is often possible to eliminate the integrator.

All modems including DM overload when the amplitude of the modulating baseband signal exceeds the range of the active devices. However, DM exhibits an additional type of overload. This overload appears when the modulating signal changes by an amount greater than the size of a step. It depends on the slope of $f(t)$ rather than its amplitude as indicated in Fig.1.20.

1.12.3 Drawbacks of DM:

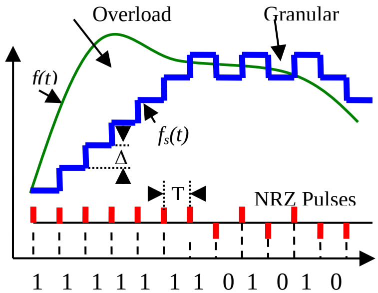

Fig.1.20: Problems of DM

Granular Noise:

When the changes in the signal $f(t)$ are less than the step size, DM no longer follows the signal and it produces a train of alternating positive and negative pulses as in Fig.1.20:

- Is similar to quantization noise in conventional PCM.
- When $f(t)$ has constant amplitude, the reconstructed signal has variations that were not present in the original signal.
- Can be reduced by decreasing the step size (high resolution).

Slope Overload Distortion:

- If analogue signal changes at a faster rate, $f(t)$ of high slope.
- Its slope is greater than DM can maintain as in Fig.1.20.
- Can be reduced by:
    - Increasing the clock frequency.
    - Increasing the step size (Low resolution)

We should compromise between these requirements, or otherwise have some control.

1.12.4 Adaptive Delta Modulation ADM

Both limitations may be overcome by adjusting the step size in accordance with the signal as in the adaptive delta modulation

ADM. Adaptive delta modulation ADM block diagram is indicated in Fig.1.21.

If $f(t)$ is constant or nearly constant, the pulses $x_c(t)$ will alternate in sign. Thus the dc value, over the time constant of LPF, is nearly zero. If $f(t)$ is increasing or decreasing rapidly, the pulses $x_c(t)$ will have the same polarity over this period. Thus the magnitude of the low pass filter output will be relatively large. The result is an increase in the gain of the variable gain amplifier. Thus in turn reduces the time of significant slope overload.

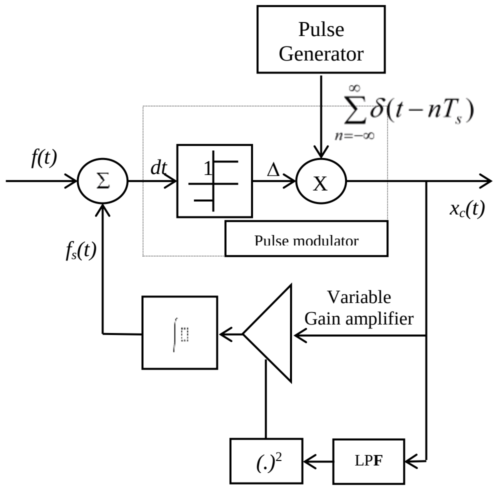

Fig.1.21: Adaptive Delta Modulation

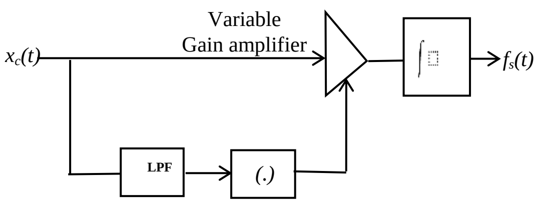

Fig.1.22 Adaptive Delta

- If $f(t)$ is relatively constant, pulses $x_c(t)$ will alternate sign.
  - dc value of LPF output is nearly zero (minimum step size)
- If $f(t)$ increases or decreases rapidly, $x_c(t)$ will have the same polarity over that period.
  - dc value of LPF output is large (large step size).

The receiver of ADM should be adaptive also as indicated in Fig.1.22.

1.13 Pulse Code Modulation PCM

Pulse modulation techniques: PDM and PPM are usually used in special purposes often military. Whereas PAM could be used as an intermediate stage within PCM, PSK, and QAM digital formats. Generally, sampling allows us to time division multiplex a number of messages whereas quantization could be used to reduce the effect of noise. The combined operations of sampling and quantizing generate a quantized PAM signal.

- The quantized sample value may be transmitted directly as in DM and ADM techniques.
- Alternatively each quantized sample is mapped to a code number in order to transmit it rather than the sample value. This is what is so called pulse code modulation "PCM".

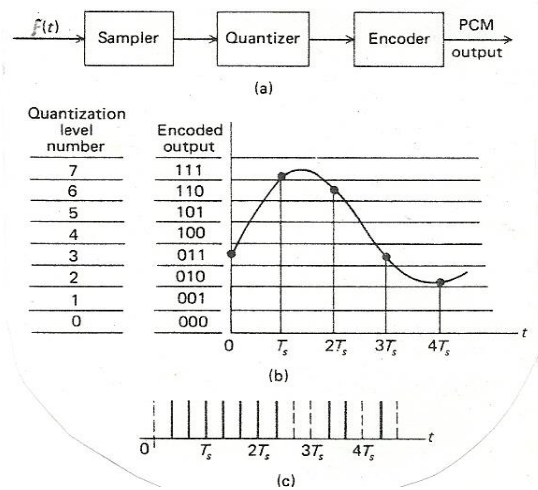

Fig.1.23: Generation of PCM:
(a) Modulator,
(b) Quantization-encoding,
(c) PCM Output

Generation of PCM formats consists of three processes as shown in Fig.1.23.a. The message signal $f(t)$ is first sampled. The sample values are then quantized as indicated in Fig.1.23.b. The quantization level of each sample is the transmitted quantity instead of the sample value. This quantization level is encoded into binary sequence as indicated in Fig.1.23.c.

In general, any quantized signal sample may be coded into a group of *m* pulses, each with *n* possible amplitude, so that the total number of possible quantization levels *M* resulting from *m* pulses is given by:

$$M = n^m \tag{1.24}$$

A binary code is just a special case for PCM encoding where *n* equals 2, hence the number of quantization levels *M* is related to the number of bits per sample *m* as follows:

$$M = 2^m \tag{1.25}$$

If the maximum frequency of the original signal is f_m so that the sampling rate is $2f_m$, then $2mf_m$ binary pulses must be transmitted per second. So, the maximum width of each binary pulse is given by:

$$\tau_{\max} = \frac{1}{2mf_m} \tag{1.26}$$

Generally, the bandwidth required for transmission of pulse train is inversely proportional to its width and depends on its shape.

Therefore, bandwidth is roughly given by the reciprocal of the time slot. This reciprocal time-bandwidth relation represents a very useful rule of thumb. Hence, the minimum bandwidth of PCM system is roughly estimated as:

$$B = 2mf_m \tag{1.27}$$

Fig.1.24 shows an example of 10 channels PAM, Quantized, and PCM speech signals using 8 levels quantizer. If the pulses were allowed to occupy the full time slots as indicated. The bandwidths required at the three points shown would be approximately as:

- PAM signal bandwidth is 1/12.5 μsec = 80 kHz.
- Quantized PAM signal bandwidth is 80 kHz.
- PCM signal bandwidth is 1/ 4.2 μsec = 240 kHz.

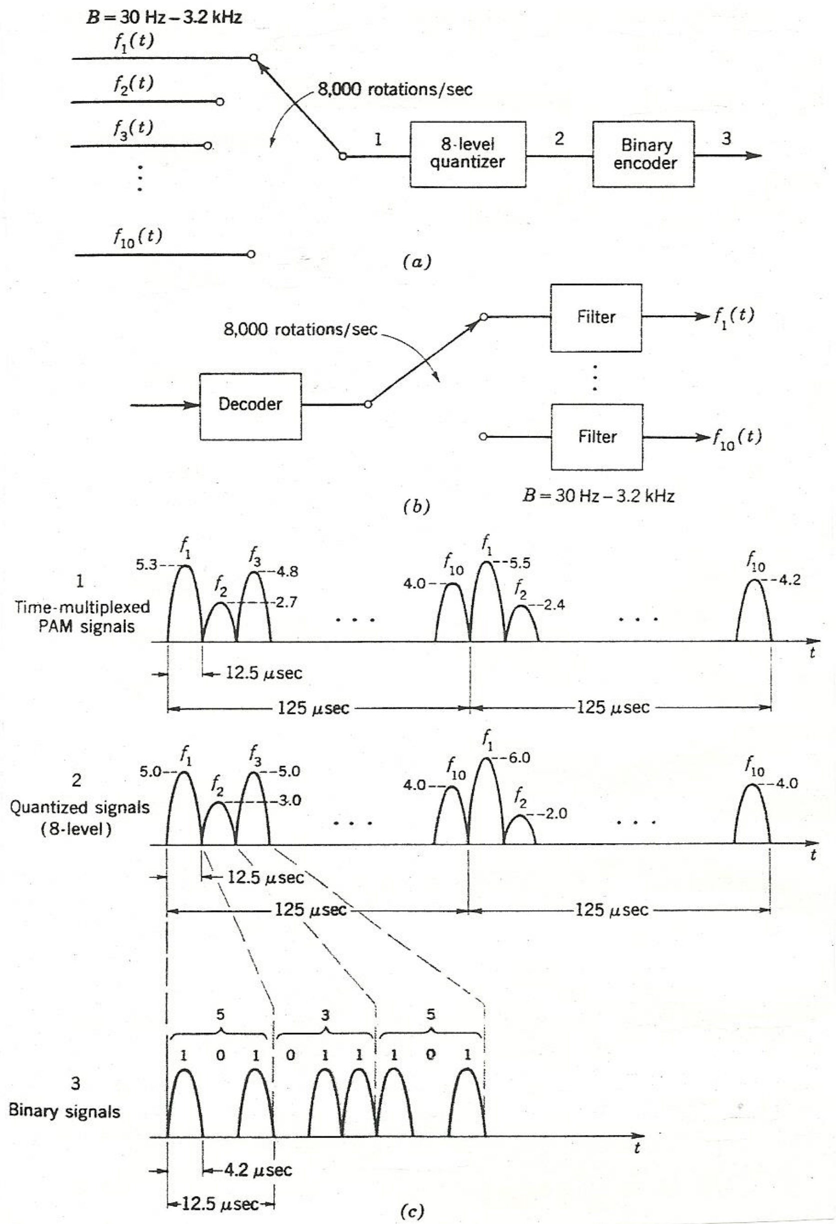

Fig.1.24: 10-Channels PCM: (a) Transmitter, (b) Receiver, (c) Signal Shapes

Exercise 1.4: What is the difference between:

- Unipolar and bipolar quantization?
- Midrise and mid-tread quantization?
- Uniform and non-uniform quantization?

Exercise 1.5:

Comment on the difference between the block diagram of PCM systems given in Fig.1.23 and the following Fig.1.25?

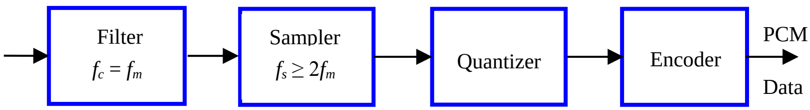

Fig.1.25 General Block Diagram of PCM System

Exercise 1.6:

Illustrate the function of each block of above PCM system.

Exercise 1.7:

Indicate the advantages of PCM systems when compared to PAM or Delta techniques.

Exercise 1.8:

Consider the design of PCM system for a bipolar analogue signal assuming the use of the symmetric folded binary code. Assume the maximum step size is 1V and the maximum peak value of the signal is 3V. [Hint: Assume sign-magnitude codes and reconstruct Fig.1.23].

Example 1.2:

For unipolar analogue signal its output maximum could not exceed v_{max} =10V. Assume you are going to encode it into n equals 8 bits. Estimate the following:

- The normalized value of each step.
- The value of voltage for each step.

- The maximum normalized value for the signal.
- The maximum voltage value for the signal
- The maximum value of quantization error in voltage.

Answer

Normalized Step Size: $\Delta X_u = 2^{-n} = 2^{-8} = \dfrac{1}{2^8} = \dfrac{1}{256} = 0.0039$

Step Size in Voltage: $\Delta V_u = \Delta X_u . v_{\max} = 0.0039x10 = 0.039V$

Maximum Normalized Signal:

$$\Delta X_{u,\max} = 1 - \Delta X_u = 1 - 0.0039 = 0.9961$$

Maximum Signal in Voltage:

$$\Delta V_{u,\max} = 1 - \Delta V_u = 1 - 0.039 = 9.961V$$

Maximum Quantization Error:

$$e_{u,\max} = \left(\frac{\Delta V_u}{2}\right) v_{\max} = \left(\frac{0.39}{2}\right) .10 = 0.0195V$$

Exercise 1.9:

Repeat example 1.2 for the case of bipolar analogue signal assuming the maximum level for the signal is limited to 10V also?

1.13.1 Companding (Compression and Expansion)

Small signal will have a poorer signal to quantization noise ratio SQR than large signal. So it is better to have moderate levels for both low and high variations of the signal:

- Increase the low signal levels to moderate levels so that SQR could be increases.
- Reduce the high signal levels to moderate levels and hence forth increases SQR.

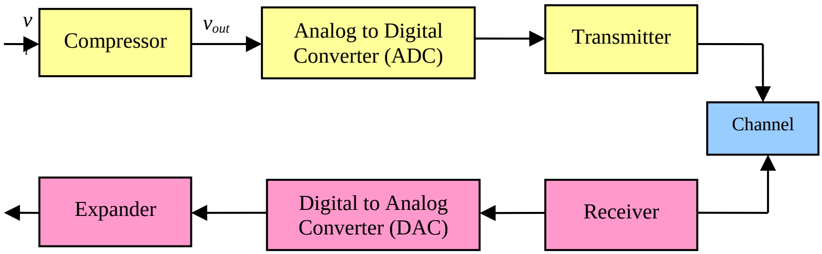

Fig.1.26 General Block Diagram of PCM System

This could be simple obtained through using what is so called as companding process. Companding is achieved by two process compression and expansion as indicated in Fig.1.26.

Analogue **compression** can be done using specially designed diodes prior to sample and hold circuit whereas **expansion** is attained with diodes after the receiver LPF.

Voice signal require a constant SQR over a wide dynamic range DR. This requires a logarithmic compression ratio. There are two methods that closely approximate a logarithm:

- μ Law Companding.
- A Law Companding.

1.13.2 μ Law Companding (USA and Japan)

The compression characteristics is given as follows where the value of μ determines the range of signal power over which SQR is relatively constant:

$$v_o = \frac{\ln\left[\left(1+\mu\left(\frac{v_i}{V_{i,\max}}\right)\right)\right]}{\ln[1+\mu]} V_{o,\max} \qquad (1.28)$$

- $V_{i,\,max}$: Max amplitude of input signal before compression.
- $V_{o,\,max}$: Max amplitude of output signal after compression.

- v_i: Amplitude of the input signal before compression ($v_i > 0$).
- v_o: Amplitude of the output signal after compression.

Whereas μ is a measure for the amount of compression as in Fig.1.27. For $DR = 40$ dB voice transmission:

- 7 bit PCM code uses $\mu = 100$.
- 8 bit PCM code uses $\mu = 255$.

On reception, received signal should be expanded to satisfy linearity for the whole compression and expansion processes.

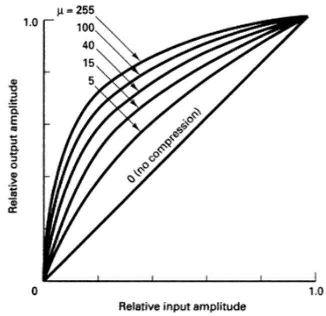

Fig.1.27: μ Law Companding

Henceforth, the received signal after expansion is calculated from the following characteristics:

$$v_{o,r} = \frac{V_{i,r,\max}}{\mu}\left[(1+\mu)\frac{v_{i,r}}{V_{o,r,\max}} - 1\right] \& v_{i,r} \geq 0 \qquad (1.29)$$

- $V_{i,,r,max}$: Max input received signal before expansion.
- $V_{o,r,,max}$: Maximum output received signal after expansion.
- $v_{i,r}$: Amplitude of the input received signal before expansion.
- $v_{o,r}$: Amplitude of the output received signal after expansion.

1.13.3 *A* Law Companding (Europe and CCITT)

The compression characteristics that is true logarithm is:

$$v_o = \frac{\dfrac{Av_i}{V_{i,\max}}}{1+\ln A} V_{o,\max} \quad , \quad 0 \leq \frac{v_i}{V_{i,\max}} \leq \frac{1}{A} \qquad (2.30)$$

$$v_o = \frac{1+\ln\left[\dfrac{Av_i}{V_{i,\max}}\right]}{1+\ln A} V_{o,\max} \quad , \quad \frac{1}{A} \leq \frac{v_i}{V_{i,\max}} \leq 1 \qquad (1.31)$$

- $V_{i,max}$: Max amplitude of the signal before compression.

- $V_{o,max}$: Max amplitude of the output signal after compression.
- v_i: Amplitude of the input signal before compression.
- v_o: Amplitude of the output signal after compression.

Whereas A is a measure of compression as shown in Fig.1.28. The optimum value for voice transmission is 87.6.

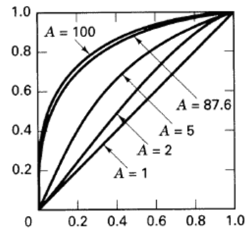

Fig.1.28: A Law Compression

1.14 North American PCM for Telephone

The Bell System in USA has introduced 24-channel PCM in the early 1960s for digital voice communication over short haul distances of 10 to 50 miles. The T1 system has found widespread adoption throughout the United States, Canada, and Japan. The 8 kHz sampling rate and the 8 bit per sample quantization form the basis for most digital PCM voice system.

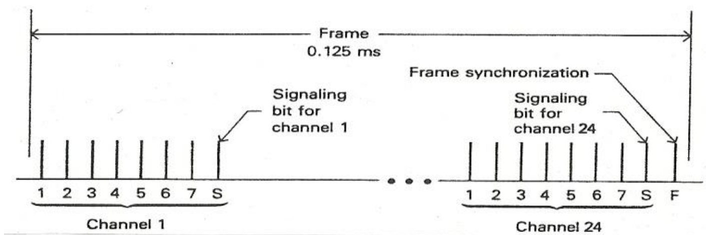

Fig.1.29: T1 PCM System for USA

The T1 system has the construction in Fig.1.29 where 24 telephone channels are time-multiplexed, sampled, and coded into PCM for carrier transmission. It used 2^7 =128 levels of quantization for each of the 24 voice channels. So, each voice signal is sampled at 8000 samples per second, and each sample is quantized into seven binary digits. An additional binary digit, known as a signaling bit, is added to the seven bits. The signaling bit is used

for establishing calls and synchronization. T1 format groups twenty-four 8-bit PCM words plus 1-bit for frame synchronization into a frame 125μsec. So, the frame consists of 24x8+1=193 bits and hence allows the transmission of 1.544 Mbps data rate.

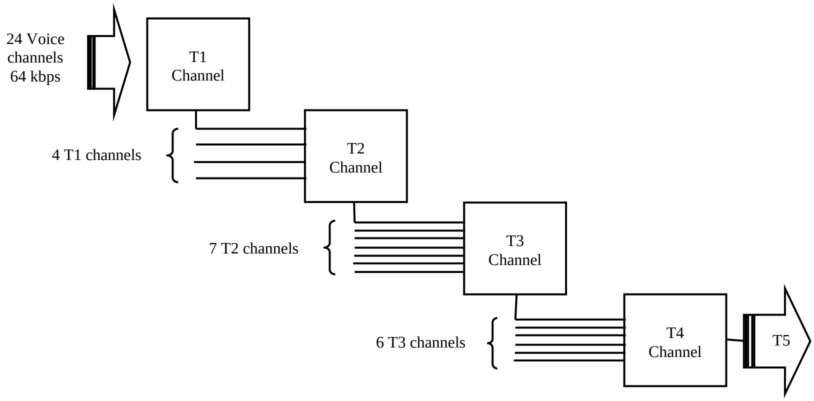

Fig.1.30: Short Haul to Long Haul PCM Systems

Exercise 1.10

Prove that the transmission data rate of T1 PCM system used in United States, Canada, and Japan is given as 1.544 Mbps.

Exercise 1.11

Estimate the number of channels and the data rate for long haul T2, T3, and T4 of Fig 2.30.

1.15 Intersymbol Interference and Wave Shaping

1.15.1 Intersymbol Interference

Main Causes:

First; when the bandwidth of the PCM **communication channel is restricted**, the transmitted waveforms will be **rounded off**, **reshaped** or **distorted** in a way entirely analogous to crosstalk in PAM.

Second; in actual practice, where digital pulses modulate a carrier for **transmission over long distances**, **pulse shaping** is a must. This is particularly true where constraints are placed on the channel bandwidth. System filtering (low and high frequency cutoffs) causes these **pulses to spread out** as they transverse the channel and they **overlap into adjacent times** as in Fig.1.31.

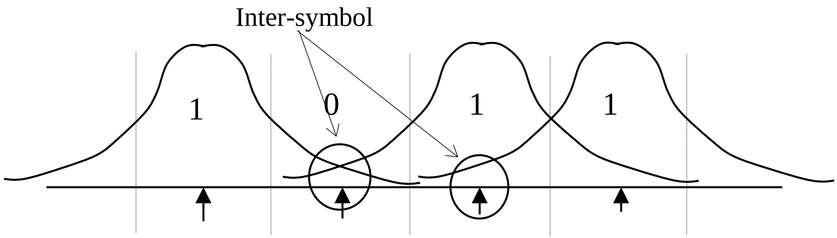

Fig.1.31: Inter-symbol Interference in Digital Transmission

Example 1.3:

Why it has been termed as crosstalk in PAM whereas in PCM it is well known as inter-symbol interference?

Answer 1.3:

- In PAM, adjacent time slots are often associated with different channels, and the term crosstalk is appropriate.
- In PCM adjacent bits are more generally symbols in the code representation of a single quantized sample, hence the term *intersymbol interference*.

Widening the transmission bandwidth as much as desired may minimize intersymbol interference. Instead, one could seek a way of purposely designing the signal wave-shapes and the transmission filters used to minimize or eliminate this interference.

1.15.2 Sampling Characteristics

One signal wave-shape producing zero intersymbol interference is the sampling function $(sin2\pi f_c t)/2\pi f_c t$ that is the

impulse response of an ideal LPF of cutoff frequency f_c. The pulse goes through zero at equally spaced intervals, the multiples of $1/2\pi f_c$ as indicated in Fig.1.32. If the sample interval is chosen $1/2\pi f_c$ adjacent pulses will not interfere as indicated in Fig.1.33.

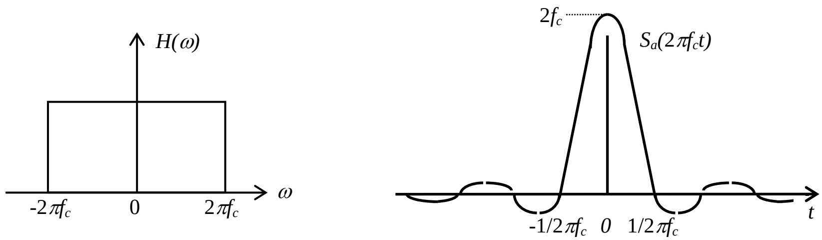

Fig.1.32: Pulse Providing Zero Intersymbol Interference

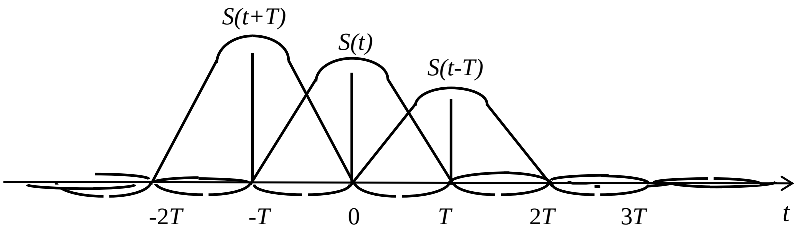

Fig.1.33: Sequence of Digital Pulses with Zero Inter-symbol Interference

However, there are practical difficulties with this particular wave-shaping since:

- It implies that the overall characteristic between the transmitter and receiver points is that of an ideal LPF. This is physically unrealizable and very difficult to approximate in practice.
- It would require extremely precise synchronization. If timing at the receiver varies somewhat from exact synchronization, the zero intersymbol interference condition disappears. In fact, the tails of all adjacent pulses may add up.
- Also some timing jitter will be present even with the most sophisticated synchronization.

Alternatively, is to modify the ideal LPF to attain a more gradual frequency cutoff, which is hence more readily realizable.

If the pulse is designed to have odd symmetry about LPF cutoff point, the resultant impulse response retains the property of having zeros at uniformly spaced interval.

1.15.3 Raised Cosine

An example is the raised cosine amplitude spectrum as in Fig.1.34 and given mathematically as:

$$\begin{aligned} H(\omega) &= \frac{1}{2}\left(1+\cos\frac{\pi\omega}{2\omega_c}\right) \qquad \& \qquad \omega \le 2\omega_c \\ &= 0 \qquad\qquad\qquad\qquad\qquad\quad elsewhere \end{aligned} \tag{1.32}$$

To show that it has an odd symmetry about ω_c, let $\omega = \omega_c + \Delta\omega$:

$$H(\omega) = \frac{1}{2}\left(1+\cos\left\{\frac{\pi}{2}+\frac{\pi}{2}\frac{\Delta\omega}{\omega_c}\right\}\right) = \frac{1}{2}\left(1-\sin\frac{\pi}{2}\frac{\Delta\omega}{\omega_c}\right) \tag{1.33}$$

Since the sine term has odd symmetry, sin -*x* = - sin *x*, so the raised cosine spectrum displays the odd symmetry as indicated in Fig.1.34. The impulse response of a filter with the characteristics shown is given by:

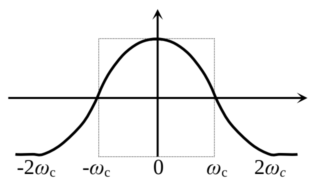

Fig.1.34: Raised Cosine

$$h(t) = \frac{\omega_c}{\pi}\frac{\sin\omega_c t}{\omega_c t}\left(\frac{\cos\omega_c t}{1-(2\omega_c t/\pi)^2}\right) \tag{1.34}$$

It has the $(\sin x)/x$ term multiplied by an additional factor decreases with increasing time. The $(\sin x)/x$ ensures zero crossing as for LPF. The additional factor reduces the tails of the pulses so that it becomes insensitive to timing jitter.

As an example, if the analog speech of 3.2 kHz is sampled at 8 kHz rate for PAM transmission, the bandwidth required for the ideal LPF approximation would be 4 kHz while for raised cosine spectrum should be 8 kHz.

So, raised cosine spectrum doubles the bandwidth required. The raised cosine spectrum is just one example of a class of spectra with odd symmetry about ω_c providing zero crossing at equally spaced sampling intervals.

1.16 Digital Signaling Formats

1.16.1 Binary Line Coding

Binary 1's and 0's may be represented in various serial signaling formats or line codes. Fig.1.35 shows the most popular:

- **Unipolar NRZ** (on-off keying): The binary 1 is represented by a high level and a binary 0 by a zero level. High level does not return to zero during the binary 1 signaling interval.
- **Polar NRZ** (simply NRZ): Binary 1's and 0's are represented by equal positive and negative levels.
- **Unipolar RZ**: The binary 1 is represented by a high level over half of the bit period and then returns to zero.
- **Bipolar**: The binary 1 is represented by alternatively positive or negative values over a half-bit period. The binary 0 is represented by a zero level.
- **Manchester** (split phase encoding): Each binary 1 is represented by a positive half-bit period pulse followed by a negative half-bit period pulse. Similarly, a binary 0 is represented by a negative half-bit period followed by a positive half-bit period.

1.16.2 Desirable Properties of Line Codes

- ***Self-synchronization:*** There is enough timing information built into the code so that bit synchronizer can be designed to extract the clock signal. So, long series of binary 1's and 0's should not cause a problem in time recovery.

- ***Error detection capability:*** The ability of addition of the channel encoders and decoders or incorporating them into the line code.
- ***Low probability of error:*** Receivers can be designed to recover the data with low probability of error when the input is corrupted by noise or intersymbol interference ISI.
- ***Suitable spectrum for channel:*** With ac coupled, the power spectral density of line code should be negligible at frequencies near zero. Also, the signal bandwidth needs to be sufficiently small compared to channel bandwidth to minimize ISI.
- ***Transparency:*** Every possible sequence of data is faithfully and transparently received. A code is not transparent if some sequence will result in a loss of synchronization at the receiver. For example, the bipolar format is not transparent since a string of zeros will result in a loss of clocking signal.

Unipolar or polar signaling may have a nonzero dc level, its value depends on the exact data. On the other hand, bipolar and Manchester signaling will always have a zero dc regardless of the data being transmitted.

1.16.3 Power Spectra of Line Codes and Comparison

- **Unipolar and Polar Signaling NRZ:** The disadvantage of unipolar and polar signaling is the waste of power due to the dc level (Fig.1.36.a and b) so that dc coupled circuits are needed. The advantage is that they are easy to generate (TTL and CMOS circuits). Polar signaling requires one supply although unipolar signaling requires positive and negative supplies. Polar signaling has a superior bit error performance than other signaling methods.
- **Unipolar RZ:** In addition to the continuous spectrum, it has discrete spectral lines at the odd multiples of the bit rate. It requires twice of the bandwidth of that of the NRZ codes.

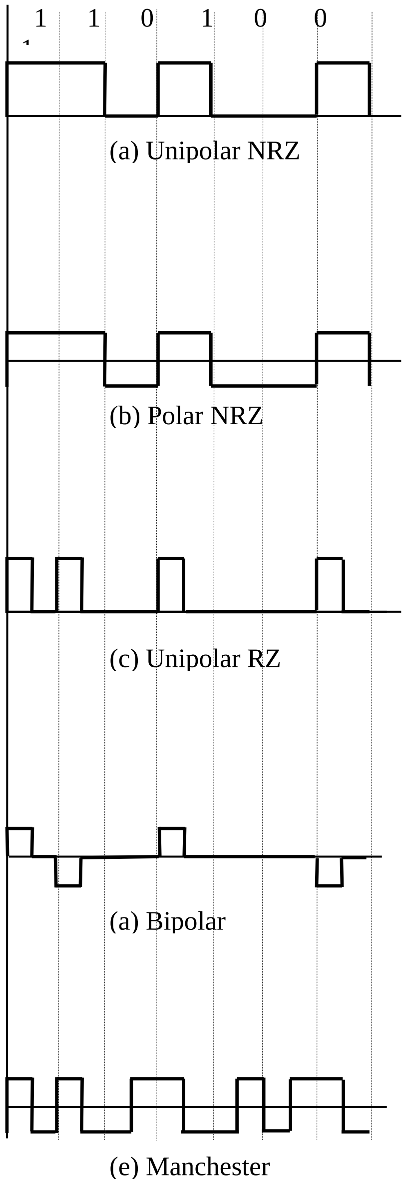

Fig.1.35: Binary Signaling Formats

Fig.1.36: Typical *psd* for Line Codes

- **Bipolar:** There is neither DC nor discrete spectral lines. Its bandwidth is that of NRZ.
- **Manchester:** There is neither DC nor discrete lines. Its bandwidth is twice that of NRZ.

1.17 Spectral Efficiency

The spectral efficiency of a digital signal is given by the number of bits per second of data that can be supported by each hertz of bandwidth. So, if the *R* is the bit rate and *B* is the signal bandwidth:

$$\eta = \frac{R}{B} \quad bits/\sec/Hz \tag{1.35}$$

The spectral efficiency can be easily evaluated from the power spectral density of Fig.1.36. Unipolar, polar, and bipolar signals are twice as efficient as RZ or Manchester signals.

However much greater efficiency can be achieved with multilevel signaling whereas the theoretical limit is determined by Shannon’s capacity formula:

$$\eta = \frac{C}{B} = \log\left[1 + \left(\frac{S}{N}\right)\right] \tag{1.36}$$

Exercise 1.11: Calculate the spectral efficiency of line codes assuming the first null bandwidth of the power spectral density.

1.18 Differential Coding

When serial data are passed through many circuits along a communication channel, the waveform is often unintentionally inverted (i.e., data complemented). This can happen in a twisted-pair transmission line channel just by switching the two leads at a connection point when a line code such as polar signaling is used.

For this reason differential coding is often employed as indicated in Fig.1.37. The encoded differential data are generated mathematically by:

$$e_n = d_n \oplus e_{n-1} \qquad (1.37)$$

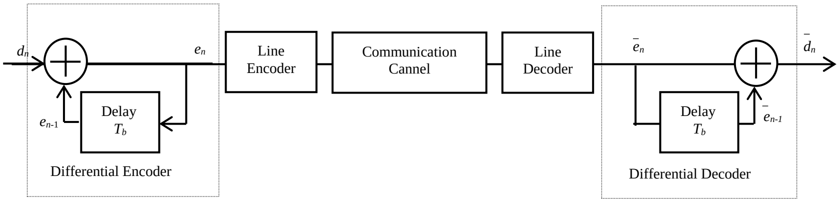

Fig.1.37: Differential Coding System

Where $\oplus$ is the modulo 2 adder, or an exclusive OR gate operation. The received encoded data are:

$$\bar{d}_n = \bar{e}_n \oplus \bar{e}_{n-1} \qquad (1.38)$$

Each digit in the encoded sequence is obtained by comparing the present input bit with the past encoded bit. A binary 1 is encoded if the present bit and past encoded bit are of opposite state, and a binary 0 is encoded if the states are the same.

1.19 Synchronous and Asynchronous Lines

Time division multiplexing (TDM) is the time interleaving of samples from several sources so that the information from these sources can be transmitted serially over a single communication channels.

It is knows that intersymbol interference (ISI) due to poor channel filtering would cause samples from one channel to appear on another channel even though perfect synchronization is maintained (crosstalk).

In the synchronous system, each device is designed so that internal clock is relatively stable for a long period of time, and it is

synchronized to a system master clock. Therefore, each bit of data is clocked in synchronism with the master clock.

For example, US telephone industry plans to synchronize all of the digital lines with a master clock located in Hillsboro. It requires a higher levels of synchronization to allow the receiver to determine the beginning and end of each block of data. The synchronizing signal may be:

- provided by a separate clocking line, or
- embedded in the data signal (using Manchester line codes)

In asynchronous systems, the timing is precise only for the bits within each character (or word). It is called ***start-stop*** signaling because each character consists of a ***start bit*** that starts the receiver clock and concludes with one or two ***stop bits*** that terminate the clocking. The receiver clock is started aperiodically and no synchronization with a master clock is required.

For example, keyboard terminals are asynchronous sources. The complete character consists 10 10 bits: one start bit, 7-bit of the ASCII code, one parity bit, and one stop bit (for $R>300$ bits/sec). In asynchronous TDM, different sources are multiplexed on a character-interleaved basis instead of interleaving on a bit-by-bit basis.

Synchronous transmission system is more efficient because start and stop bits are not required. However, synchronous mode requires that the clocking signal be passed along with the data and the receiver synchronize to the clocking signal.

Multiplexers may be classified according synchronization:

1. TDM to synchronous lines.
2. TDM to quasi-synchronous lines: Individual clocks of the input data sources are not exactly synchronized in frequency. So there will be some variation in the bit rates from different sources. Sometimes, the clock rates

are not related by a rational number. This requires stuff bits, which are dummy bits; 1's, 0's, or some alternating pattern, to compensate for the difference sources rates.

3. TDM to asynchronous lines: It produces:

   - High-speed asynchronous output (no stuff bits is required).
   - High-speed synchronous output (staff bits is required).